

Attorney Docket No: 23540-07445/US

Client Ref: 2001-072-2

USSN: 09/955,663

**AMENDMENTS TO THE CLAIMS****1. (Canceled)**

**2. (Currently amended):** A method for estimating the precision of a set of gene expression data comprising observed intensity measurements taken from at least one a-gene expression microarray experiment, said method comprising:

inputting the set of gene expression data:

- (a) identifying a first subset set of low-level data observed intensity measurements from the set of gene expression data, wherein said subset of low level observed intensity measurements have values less than an observed intensity measurement cutoff determined by a thresholding algorithm;
- (b) estimating a standard deviation,  $\sigma_e$ , of an additive error component,  $\varepsilon$  of the set of gene expression data using either at least one negative control or using the first subset of low level observed intensity measurements;
- (c) estimating a mean background intensity measurement parameter,  $\alpha$ , using either at least one negative control or using the first subset of low level observed intensity measurements;
- (d) identifying a second subset set of replicated high-level data observed intensity measurements from the set of gene expression data, wherein the second subset comprises several highest intensity measurements of the set or the variance of logarithms of the second subset is approximately constant;
- (e) estimating a standard deviation,  $\sigma_\eta$ , of a proportional error component,  $\eta$ , of the set of gene expression data using from the standard deviation of the logarithm of the second subset of replicated high-level observed intensity measurements; and data set;
- (f) measuring a signal,  $y$ , wherein said signal indicates an amount of a biological molecule; and
- (g) using the estimates of  $\sigma_e$  and  $\sigma_\eta$  and  $\alpha$  for estimating a variance,  $Var$ , of the observed intensity measurements,  $y$ , measured signal as

Attorney Docket No: 23540-07445/US  
 Client Ref: 2001-072-2  
 USSN: 09/955,663

$$\text{Var}\{y\} = \mu^2 e^{\sigma_a^2} (e^{\sigma_a^2} - 1) + \sigma_e^2$$

wherein  $\mu$  is an expression level of a biological molecule in arbitrary units ~~the amount of the biological molecule~~ and  $y = \alpha + \mu e^{\eta} + \varepsilon$ ; and

using the variance to estimate the precision of the gene expression data set.

~~and wherein said identifying step (a) comprises the use of a thresholding algorithm to establish a cutoff, and the set of low-level data consists of those data with values less than the cutoff.~~

**3. (Currently amended):** The method of claim 2, wherein the thresholding algorithm comprises the steps of:

- (a) identifying  $A_N$ , an initial set of low-level data measurements consisting of  $q$  percent of the total number of data points having the lowest measurement values,  $A_N = \{x_1, x_2, \dots, x_{nq}\}$ ;
- (b) calculating a mean and a standard deviation of the initial set,  $A_N$ ;
- (c) calculating a cutoff point,  $\mu_N = \text{mean} + (c \times \text{the standard deviation})$ , wherein  $c$  is the number of standard deviations or mean absolute deviations above the mean or median and  $c$  is set at 2, 2.5 or 3 for each iteration of the thresholding algorithm;
- (d) defining a new set,  $A_{N+1} = \{x_j < \mu_N\}$ ;
- (e) calculating a mean and standard deviation of the new set; and
- (f) repeating steps (c) and (d) using the mean and standard deviation of the new set until  $n_k = n_{k-1}$  ~~the algorithm converges.~~

**4. (Currently amended):** The method of claim 2, wherein the thresholding algorithm comprises the steps of:

- (a) identifying  $A_N$ , an initial set, of low-level data consisting of  $q$  percent of the total number of data points having the lowest measurement values,  $A_N = \{x_1, x_2, \dots, x_{nq}\}$ ;

Attorney Docket No: 23540-07445/US

Client Ref: 2001-072-2

USSN: 09/955,663

- (b) calculating a median of the initial set,  $\underline{A}_N$ ,  $m_0 = \text{median} \{x_j\}_{j=1}^{n_0}$  and a median of the absolute deviations about the median,  $MAD_0 = \text{median} \{x_j - m_0\}_{j=1}^{n_0}$ ;
- (c) calculating a cutoff point,  $u_0 = MAD_0 + (c \times s_0)$ , wherein  $s_0 = MAD_0/0.675$  and  $c$  is the number of standard deviations or mean absolute deviations above the mean or median and  $c$  is set at 2, 2.5 or 3 for each iteration of the thresholding algorithm;
- (d) defining a new set,  $A_{N+1} = \{x_j < u_N\}$ ;
- (e) calculating a median and a median of the absolute deviations about the median of the new set; and
- (f) repeating steps (c) and (d) using the median and the median of the absolute deviations about the median of the new set until  $n_k = n_{k-1}$  ~~the algorithm converges~~.

**5. (Original):** The method of claim 2, wherein the mean of the low-level data measurements is used as the estimate of the background parameter,  $\alpha$ .

**6. (Previously presented):** The method of claim 2, wherein the standard deviation of the low-level data measurements is used as the estimate of the parameter  $\sigma_x$ .

**7. (Previously presented):** The method of claim 2, wherein a mean of negative control data is used as the estimate of the background parameter,  $\alpha$ .

**8. (Previously presented):** The method of claim 2, wherein the biological molecule is a nucleic acid.

**9. (Original):** The method of claim 8, wherein the nucleic acid is mRNA.

**10. (Original):** The method of claim 8, wherein the biological molecule is DNA.

**11. (Original):** The method of claim 10, wherein the DNA is cDNA.

**12. (Original):** The method of claim 10, wherein the DNA is genomic.

**13. (Canceled)**

**14. (Canceled)**